WAIS Digit Span–Based Indicators of Malingered Neurocognitive Dysfunction

Classification Accuracy in Traumatic Brain Injury

Matthew T. Heinly
Kevin W. Greve
Kevin J. Bianchini
Jeffrey M. Love
University of New Orleans & Jefferson Neurobehavioral Group

Adrianne Brennan
Jefferson Neurobehavioral Group

*The present study determined specificity and sensitivity to malingered neurocognitive dysfunction (MND) in traumatic brain injury (TBI) for several Wechsler Adult Intelligence Scale (WAIS) Digit Span scores. TBI patients (*n = *344) were categorized into one of five groups: no incentive, incentive only, suspect, probable MND, and definite MND. Performance of 1,063 nonincentive patients (e.g., cerebrovascular accident, memory disorder) was also examined. Digit Span scores included reliable digit span, maximum span forward both trials correct, maximum span forward, combined maximum forward and backward span, Digit Span scaled score, maximum span backward both trials correct, and maximum span backward. In TBI, sensitivity to MND ranged from 15% to greater than 30% at specificities of 92% to 98%. Patient groups with documented brain pathology had higher false-positive error rates. These results replicate previous known-groups malingering studies and provide valuable data supporting the WAIS Digit Span scores in detection and diagnosis of malingering.*

Keywords: digit span; reliable digit span; malingering; response bias; traumatic brain injury; neuropsychological assessment

Psychological and neuropsychological evaluations done in the context of incentive (i.e., worker's compensation and personal injury litigation cases) necessitate assessment of patient performance validity. One approach to evaluating the validity of a patient's performance and detecting malingering of cognitive and perceptual symptoms is the stand-alone, forced-choice symptom validity test (SVT; for a review of SVTs, see Bianchini, Mathias, & Greve, 2001). An alternative to the stand-alone SVT is the use of internal or embedded internal validity indicators derived from standard neuropsychological instruments. There are three reasons that internal validity indicators have great utility: (a) They enhance the sensitivity of the entire malingering battery without increasing the time

Cora Joffe and Nora Vitone assisted with data collection and entry; their efforts are much appreciated. Jeffrey M. Love is now at Pennsylvania State University and Adrianne Brennan is now at Louisiana State University. Portions of this research were presented at the Annual Meeting of the National Academy of Neuropsychology, November 2004, Seattle, Washington. Requests for reprints should be sent to Kevin W. Greve, Department of Psychology, University of New Orleans—Lakefront, New Orleans, LA 70148; phone: 504-280-6185; fax: 504-280-6049; e-mail: kgreve@uno.edu.

Assessment, Volume 12, No. 4, December 2005 429-444
DOI: 10.1177/1073191105281099

required for the assessment, (b) they can provide information about the validity of performance on specific tests (Mathias, Greve, Bianchini, Houston, & Crouch, 2002; Meyers & Diep, 2000; Meyers & Volbrecht, 1998), and (c) they may be less likely to be coached than SVTs (Mathias et al., 2002). Numerous indicators have been reported during the past 10 years (e.g., Bernard, McGrath, & Houston, 1996; Greiffenstein, Baker, & Gola, 1994; Greiffenstein, Gola, & Baker, 1995; Millis, Putnam, Adams, & Ricker, 1995; Mittenberg et al., 2001, Mittenberg, Theroux-Fichera, Zielinski, & Heilbronner, 1995; Suhr & Boyer, 1999).

Digit Span, which is administered as part of several commonly used clinical instruments (e.g., Wechsler Intelligence and Memory Scales, see Wechsler, 1981, 1987, 1997a, 1997b; Stanford-Binet, see Thorndike, Hagen, & Sattler, 1986; Neuropsychological Assessment Battery, see Stern & White, 2003; Cognistat, see Kiernan, Mueller, & Langston, 1983), may be one of the most ubiquitous psychological tests. Superficially, it appears to be a memory test that would be difficult for someone with brain pathology, and one would hypothesize that a patient attempting to appear impaired may perform poorly on it (Meyers & Volbrecht, 1998). In actuality, Digit Span performance is fairly well preserved in persons with brain dysfunction including those with amnestic disorders (Butters & Cermak, 1980; Greiffenstein et al., 1994; Iverson & Franzen, 1996; Iverson & Tulsky, 2003). Thus, Digit Span may be particularly well suited for use in malingering detection.

The utility of the Digit Span as a malingering detection method has been investigated in a number of studies using a variety of experimental designs. Iverson and Franzen (1994) examined the accuracy of several Digit Span–based scores (scaled score, raw digit span forward and backward) using a mixed-simulator/clinical-specificity design study. Digit Span was administered to college students and federal inmates asked to fake memory problems, normal controls (both students and inmates), and moderate/severe traumatic brain injury (TBI) patients who were nonlitigating and considered to be giving good effort as indicated by the 21-item test. Scores associated with a 0% false-positive error rate in both the normal controls and TBI patients detected between 52.5% and 75% of the simulated malingerers. When Iverson and Franzen (1996) extended this research to simulators who were asked to simulate memory impairment and were provided incentive for their performance, they detected between 57% and 77% while maintaining a 5% false-positive error rate in noncompensation-seeking psychiatric and memory-impaired patients. Iverson and Tulsky (2003) found that the cutoffs reported by Iverson and Franzen (1994, 1996) were associated with a false-positive rate of approximately 5% in both the normal and clinical groups of the Wechsler Adult Intelligence Scale—III (WAIS-III; Wechsler, 1997a) standardization sample. Using a known-groups design, Babikian, Boone, Lu, and Arnold (in press) found that the Iverson and Franzen cutoff for the Digit Span scaled score detected 18% of probable malingerers (diagnosed according to the criteria proposed by Slick, Sherman, & Iverson, 1999) with a false-positive error rate of 2%. Other Digit Span scores examined by this group were associated with sensitivities ranging from 20% to 30% with a false-positive error rate of 5% or less.

Also derived from Digit Span, the Reliable Digit Span test (RDS; Greiffenstein et al., 1994) may be one of the most well-validated clinical malingering indicators (e.g., Babikian et al., in press; Duncan & Ausborn, 2002; Etherton, Bianchini, Ciota, & Greve, 2005; Etherton, Bianchini, Greve, & Heinly, 2005; Greiffenstein et al., 1994; Greiffenstein et al., 1995; Inman & Berry, 2002; Larrabee, 2003a; Mathias et al., 2002; Meyers & Volbrecht, 1998; Strauss et al., 2002). RDS is calculated by summing the last forward and backward digit strings in which both trials were completed without error. Although the literature is variable, RDS scores of 7 or less have generally been associated with a false-positive error rate of 10% or less in nonmalingering brain-injured patients, those in clinical pain, criminal forensic populations, and healthy samples. RDS scores of 6 or less are rare or nonexistent in nonmalingering patients with a range of clinical conditions and therefore indicate poor effort and/or the presence of negative response bias. Overall, these findings demonstrate that scores derived from Digit Span are not just tapping into working memory but are also sensitive and specific measures of effort.

Purpose

Greve and Bianchini (2004) emphasized the importance of studying malingering indicators using a known-groups methodology and, if possible, including a range of nonmalingering clinical patients to establish the limits of specificity. Most research involving Digit Span indicators has used the known-groups design, which is considered the strongest approach for validating malingering indicators (Greve & Bianchini, 2004; Larrabee, 2005; Rogers, 1997). Although the RDS is a well-validated indicator whose classification accuracy is very good, there are some shortcomings in the extant Digit Span/malingering literature in general, which, if addressed, would greatly enhance the clinical utility of scores derived from the test. In particular, classification accuracy has generally been reported for a very limited range of RDS scores (exceptions include Etherton, Bianchini, Ciota, et al., 2005; Etherton, Bianchini, Greve, et al., 2005; and Mathias et al., 2002)

and other Digit Span scores. Unlike other indicators (e.g., the Millis and Mittenberg formulas), Digit Span scores have a relatively limited range to begin with. Reporting classification accuracy data for the full range is easily done and would allow the clinician to better interpret the scores clinically. Also, existing research has generally not provided data for comparisons within diagnostic groups. For example, in Mathias et al. (2002), the malingering group was predominantly mild TBI, whereas the nonmalingering group was predominantly moderate/severe TBI. In Babikian et al. (in press), neither the malingering nor nonmalingering groups were even limited to TBI but also contained a number of neurological and psychiatric diagnoses.

Thus, the purpose of the present study was to further examine the classification accuracy of RDS and other Digit Span scores in TBI, thereby extending the existing literature in the following ways: (a) Indices of classification accuracy are provided for all the Digit Span scores previously reported plus additional scores, thereby allowing comparisons of accuracy across scores; (b) classification accuracy data are provided for a broad range of scores rather than preselected cutoffs; (c) mild TBI and moderate/severe TBI are examined separately; (d) malingering is operationalized using the Slick et al. (1999) criteria; and (e) data for a large sample of general clinical patients without incentive (and some presented separately) are included to provide information on the effects of different conditions, thus further clarifying issues related to specificity. The results of the study are presented in frequency tables that can be easily referenced in clinical practice.

METHOD

Participants

TBI. This group consisted of 344 persons who were referred for neuropsychological evaluation after suffering an apparent TBI. Patients were considered to have suffered a mild TBI if they met the criteria set by the Mild Traumatic Brain Injury Committee of the Head Injury Interdisciplinary Special Interest Group of the American Congress of Rehabilitation Medicine (1993): (a) posttraumatic amnesia not greater than 24 hours; (b) after 30 minutes, an initial Glasgow Coma Scale of 13 to 15; and (c) loss of consciousness of approximately 30 minutes or less. Any patient who did not meet all of the mild TBI criteria and/or who had positive neuroradiological findings (e.g., skull fracture, hemorrhage, hematoma) or focal neurological signs were coded as moderate/severe TBI for purposes of this study. The mild TBI criteria were used in this study to separate mild from moderate/severe cases, not to rule in or rule out an actual TBI per se. For purposes of this study, a patient was placed in the TBI group if they presented with or claimed to have had a head injury. They were then classified as moderate/severe, if appropriate, as described above. The details of their head injury claim were examined as part of the malingering classification process (see below). Thus, the mild TBI group contained only persons who clearly had *no worse* than a mild TBI and might have had no TBI. One hundred ninety-six were placed in the mild TBI group. One hundred forty-eight were classified as moderate/severe. Overall, the TBI patients averaged 21.16 (SD = 26.02) months between the date of their injury and the date of their evaluation. There was not a significant difference in this latency between the mild and moderate/severe TBI groups, $F(1, 316) = 0.09$, $\eta^2 = .00$. The latency between injury and evaluation could not be found for 27 (8%) of the TBI patients. See Table 1 for descriptive characteristics of these groups.

General clinical sample. This group consisted of 1,063 persons seen in the course of general neuropsychological practice, including persons with a range of neurological and psychiatric diagnoses and student controls. Non-TBI patients were excluded if they were seen in a compensation-seeking context. Thus, this group is without known external incentive to perform poorly. See Table 1 for the diagnostic breakdown and demographic characteristics of this sample.

Malingering Classification

Patients were categorized on the basis of the Slick et al. (1999) criteria for malingered neurocognitive dysfunction (MND) using a diagnostic decision tree such as that presented by Millis (2004). In determining the presence of MND, the case must be evaluated on the basis of four criteria:

A. the presence of substantial external incentive,
B. evidence from neuropsychological testing,
C. evidence from self-report, and
D. behaviors meeting the necessary B and C criteria are not fully accounted for by psychiatric, neurological, or developmental factors.

Using this system, all diagnoses of malingering require the presence of external incentive (Criterion A) plus Criterion B and/or C evidence as noted below. Criterion B behaviors are sufficient for a diagnosis of malingering on their own.

Evidence from Criterion B may include discrepancies between test data and known patterns of brain functioning (B3), behavioral observations (B4), information from collaterals (B5), and documented history (B6). However,

TABLE 1
Descriptive Characteristics of the Current Sample by Diagnosis

Diagnosis	*Sample Size* n *(%)*	*Gender* % F	*Age* M *(SD)*	*Education* M *(SD)*
Mild TBI	196 (57.0)	36	41.2 (12.5)	12.3 (3.0)
Moderate/severe TBI	148 (43.0)	21	37.4 (15.1)	12.0 (3.2)
Total TBI	344 (24.4)	30	39.6 (13.8)	12.1 (3.1)
Other diagnoses				
CVA	517 (48.6)	48	66.2 (13.3)	12.0 (3.4)
Memory disorder	228 (21.4)	56	74.1 (8.8)	11.6 (3.9)
Psychiatric disorder	128 (12.0)	48	53.7 (20.5)	12.5 (3.5)
Tumor	45 (0.04)	29	52.9 (14.5)	14.0 (3.0)
Encephalopathy	41 (0.04)	39	65.3 (13.2)	11.9 (4.0)
Substance abuse	39 (0.04)	21	43.1 (17.1)	12.8 (2.6)
Infection	14 (0.01)	43	51.4 (21.2)	12.7 (3.9)
Seizure	14 (0.01)	36	47.1 (16.3)	11.1 (2.5)
Multiple sclerosis	10 (0.01)	30	44.2 (7.7)	13.2 (3.0)
Parkinson's disease	9 (0.008)	44	71.9 (6.6)	12.4 (2.2)
General medical problems	6 (0.006)	67	74.2 (6.6)	12.5 (3.3)
Lupus	4 (0.004)	100	57.8 (18.2)	15.3 (2.6)
Nonstudent control	4 (0.004)	25	57.0 (19.7)	13.3 (1.5)
Academic problems	2 (0.002)	50	17.5 (0.7)	12.5 (0.7)
Cerebral palsy	1 (0.001)	100	42.0 (0.0)	15.0 (0.0)
Learning disability	1 (0.001)	100	26.0 (0.0)	2.0 (0.0)
Total other diagnosis	1,063 (75.6)	47	64.2 (16.3)	12.1 (3.5)
Total sample	1,407 (100)	43	58.1 (18.9)	12.1 (3.4)

NOTE: TBI = traumatic brain injury; CVA = cardiovascular accident.

the most powerful Criterion B evidence is documentation of a negative response bias on the basis of performance on an SVT (e.g., Portland Digit Recognition Test [PDRT] or Test of Memory Malingering [TOMM]; see Bianchini et al., 2001, for others). Performance on a forced-choice measure can indicate either definite response bias (B1: obtained score is significantly below chance at $\alpha < .05$, two-tailed) or probable response bias (B2: obtained score on a well-validated measure of response bias is in a range consistent with exaggeration or feigning). Other malingering tests and indices from standard clinical measures can also meet B2. Criterion C behaviors include discrepancies between self-report and documented history (C1), known patterns of brain functioning (C2), behavioral observations (C3), and information from collaterals (C4). This criterion (C5) includes evidence of exaggeration or fabrication of psychological symptoms on self-report measures with well-validated validity scales (e.g., Minnesota Multiphasic Personality Inventory—2 [MMPI-2]). Criterion C indicators are considered evidence of possible malingering but are insufficient on their own for a diagnosis of malingering. For purposes of this study, Criterion B2 could be met on the basis of a positive finding on either the PDRT (Binder, 1993) or TOMM (Tombaugh, 1996) or by two or more positive findings on well-validated clinical indicators. Clinical indicators included the Millis formula for the California Verbal Learning Test (Millis et al., 1995), and the Suhr formula (Suhr & Boyer, 1999) and unique responses (Greve, Bianchini, Mathias, Houston, & Crouch, 2002) from the Wisconsin Card Sorting Test (Heaton, Chelune, Talley, Kay, & Curtiss, 1993). Criterion C5 could be met on the basis of MMPI-2 *T*-scores greater than 100 for the F, Fb, or Fp scales (Butcher, Dahlstrom, Graham, Tellegren, & Kaemmer, 1989) or an MMPI-2 Fake Bad Scale raw score greater than 30 (Larrabee, 2003b). See Table 2 for the cutoffs related to these indicators.

A B1 finding is sufficient for a diagnosis of definite MND. A diagnosis of probable MND can be made with two types of Criterion B evidence or one type of Criterion B evidence and one or more types of Criterion C evidence. Criterion C evidence is not sufficient for a diagnosis in the absence of Criterion B evidence. Possible MND is diagnosed when the criteria for probable MND have been met but Criterion D factors are present.

The entire general clinical sample lacked substantial external incentive and thus failed to meet Criterion A. These participants were placed in the no-incentive group. The TBI patients were subdivided into the following five groups (see Table 3 for the details of group assignment including breakdown by severity level):

TABLE 2
Cutoffs for Tests and Indicators Used to Meet the Psychometric Criteria (B2 and C6) Used in Diagnosing Malingered Neurocognitive Dysfunction

Indicator/Test	*Cutoff*	*Reference for Cutoff*
Portland Digit Recognition Test	Easy < 19, hard < 18, total < 39	Binder, 1993
Test of Memory Malingering	< 45 on trial 2 or retention	Tombaugh, 1996
Millis Formula (California Verbal Learning Test)	< 0	Millis et al., 1995
Suhr Formula (Wisconsin Card Sorting Test)	> 3.68	Greve et al., 2002
Unique Responses (Wisconsin Card Sorting Test)	> 1	
Minnesota Multiphasic Personality Inventory—2 Validity Scales	F, Fb, Fp *T*-scores > 100; Fake Bad Scale raw > 30	Butcher et al., 1989; Larrabee, 2003b

TABLE 3
Incentive/Malingering Status

	No Incentive n *(%)*	*Incentive Only* n *(%)*	*Suspect* n *(%)*	*Probable MND* n *(%)*	*Definite MND* n *(%)*
Mild TBI	15 (8)	62 (32)	71 (36)	36 (18)	12 (6)
Moderate/severe TBI	30 (20)	39 (26)	56 (38)	17 (12)	6 (4)
Total TBI	45 (13)	101 (29)	127 (37)	53 (15)	18 (5)

NOTE: MND = malingered neurocognitive dysfunction; TBI = traumatic brain injury.

1. No incentive ($n = 45$): These patients had no evidence of external incentive (were negative on Criterion A) and therefore could not be diagnosed as malingering. Some (40%) met none of the Slick et al. (1999) criteria.
2. Incentive only ($n = 101$): Because the no-incentive TBI group was so small, it was important to have additional TBI patients who were very likely to be giving good effort so as to more clearly define the specificity of the tested indicators. The incentive-only patients had external incentive and were negative on all B and C criteria and absent any evidence of suspect effort including any positive findings on any indicator of negative response bias. Patients meeting inclusion criteria for this group but who did not complete at least one forced-choice SVT were excluded from the study.
3. Suspect ($n = 127$): This group was defined by the presence of any evidence suggesting suspect effort including meeting any one Slick et al. (1999) criterion or being positive on any single indicator of negative response bias.
4. Probable ($n = 53$): This group met the Slick et al. (1999) criteria for probable MND.
5. Definite ($n = 18$): This group met the Slick et al. (1999) criteria for definite MND.

Procedure and Variables

The WAIS (WAIS-R and WAIS-III; Wechsler, 1981, 1997a) including Digit Span was administered as part of a comprehensive neuropsychological assessment battery. Scores examined in this study included RDS, as defined in the introduction; maximum span forward with both trials correct (MSF2); maximum span forward (MSF1), the longest forward string repeated correctly; combined maximum span forward and maximum span backward (MFB); WAIS-III Digit Span scaled score; maximum span backward with both trials correct (MSB2); and maximum span backward (MSB1), the longest backward string repeated correctly.

RESULTS

Malingering Group Analyses

The Digit Span scores for the TBI sample were submitted to univariate analyses of variance (ANOVA) to examine group effects. The means for each of the five TBI groups and analysis results are presented in Table 4. Significant group effects were observed for all Digit Span scores. Overall, the no-incentive and incentive-only groups did not differ, and the MND (probable and definite) groups generally did not differ. Thus, these results show a

TABLE 4
Digit Span Scores by Malingering Status and Results of Analyses of Variance for the Traumatic Brain Injury (TBI) Sample

	No Incentive M (SD)	*Incentive Only* M (SD)	*Suspect* M (SD)	*Probable MND* M (SD)	*Definite MND* M (SD)	F	η^2
RDS	9.91[a] (2.30)	9.93[a] (2.30)	8.32[b] (2.03)	7.15[c] (2.02)	6.61[c] (2.23)	23.7	.22
MSF2	6.00[a] (1.31)	5.79[a] (1.28)	5.13[b] (1.30)	4.42[c] (1.17)	4.11[c] (1.41)	17.9	.17
MSF1	6.71[a] (1.27)	6.73[a] (1.30)	6.02[a, b] (1.36)	5.23[c] (1.30)	5.39[b, c] (1.34)	15.2	.15
MFB	11.33[a] (2.64)	11.65[a] (2.27)	9.88[b] (2.38)	8.74[b] (2.15)	9.06[b] (2.62)	18.4	.18
Digit Span	10.48[a] (3.76)	9.74[a, b] (2.88)	8.40[b] (2.60)	6.65[c] (2.34)	6.13[c] (2.80)	14.0	.20
MSB2	3.91[a] (1.40)	4.14[a] (1.33)	3.20[b] (1.16)	2.74[b, c] (1.15)	2.50[c] (1.04)	17.9	.17
MSB1	4.62[a] (1.71)	4.92[a] (1.32)	3.86[b] (1.39)	3.51[b] (1.17)	3.67[b] (1.61)	13.5	.14

NOTE: MND = malingered neurocognitive dysfunction; RDS = Reliable Digit Span; MSF2 = maximum span forward both trials correct; MSF1 = maximum span forward; MFB = combined maximum forward and backward span; MSB2 = maximum span backward both trials correct; MSB1 = maximum span backward. *F* ratios for all variables are significant at $\alpha < .001$.
a, b, c, d, e. Row means with the same letter are not significantly different from each other.

clear relationship between the amount of evidence suggestive of cognitive malingering and low scores on Digit Span.

Classification Accuracy

The relevant indices of classification accuracy (and therefore error rate, the reciprocal of accuracy) are sensitivity, specificity, and predictive power (Hennekens & Buring, 1987; see also Gouvier, Hayes, & Smiroldo, 1998, for a discussion of these factors). Sensitivity is the true positive (hit) rate for a test (e.g., the number of malingerers who had a positive test result divided by all malingerers). Specificity is the true negative rate (e.g., the number of persons not malingering who had a negative test result divided by all persons not malingering). Predictive power is an index of the confidence one can have that an individual test result is accurate (Hennekens & Buring, 1987); it is dependent on both accuracy of the test itself and the base rate of the target condition in the population of interest.

Based on the group analyses described above, the probable and definite MND groups were combined ($n = 71$) to examine sensitivity. For purposes of estimating specificity, the no-incentive and incentive-only TBI groups were examined individually and in combination ($n = 146$). Tables 5 through 9 provide percentile levels for ranges of scores on each Digit Span variable. In each table, the value is the percentage of patients performing at or below the given score. For non-MND patients, this is the false-positive error rate (specificity = 1 – false-positive error rate); for MND patients, this is the sensitivity rate. Also included in Tables 5 through 9 are data for the mild and moderate/severe TBI cases as well as several nonmalingering non-TBI subgroups and the non-TBI subgroup overall.

Table 10 lists the percentage of TBI patients by group who scored beyond the cutoff representing the approximate 5% false-positive error rate that was estimated from the performance of the entire non-MND sample (n = 1,336) and, separately, from the nonmalingering TBI sample ($n = 146$). When specificity was based on the entire nonmalingering sample, sensitivity was generally very poor (exceptions were MSF2 and the Digit Span scaled score). However, sensitivity was much improved when accuracy was based on the nonmalingering TBI sample only. Similar to the group analyses, the frequency of positive findings on Digit Span scores in TBI generally increases as a function of the amount of evidence suggestive of cognitive malingering. Note also that the false-positive error rates for the no-incentive and incentive-only TBI groups are either below or close to 5% (with the exception of the Digit Span scaled score and MSB2). These findings suggest that objective neuropathology does have an effect on scores, particularly those that include backward span.

Positive predictive power (PPP). PPP (true positives divided by the sum of true positives and false positives; see Millis & Volinsky, 2001, for an alternative formula for the calculation of PPP or posttest odds) indicates the probability that someone is malingering given a particular score. Table 11 presents PPP at five hypothetical MND base rates for each Digit Span score. Table 11 uses the same cutoffs as reported in Table 10.

It is not practical to provide tables of PPP for all score and base rate combinations. However, the information provided in Tables 5 through 9 can be used to calculate PPP for any score and base rate combination using the method described by Millis and Volinsky (2001), which is summarized below:

(continued on page 440)

TABLE 5
Cumulative Percentages of Patients With Scores At or Below the Indicated Reliable Digit Span (RDS) Score

	All TBI					*Mild TBI*		*Moderate/Severe TBI*		*Other Diagnoses*				*Total Not*
≤ RDS	*No Inc*	*Inc Only*	*Not MND*	*SUS*	*MND*	*Not MND*	*MND*	*Not MND*	*MND*	*CVA*	*Mem*	*Psyc*	*Total Other*	*MND*
n	45	101	146	127	71	77	48	69	23	517	228	128	1,063	1,336
2				0	0		0			1	1		1	1
3				1	4		6			4	4	0	3	3
4				2	10		15		0	9	11	1	7	6
5	0	0	0	9	21	0	27	0	9	14	18	2	12	11
6	2	5	4	16	39	7	46	1	26	30	32	12	26	23
7	16	12	13	33	65	17	71	9	52	44	52	28	42	38
8	36	25	28	58	80	31	83	25	74	64	70	50	62	58
9	47	50	49	75	90	55	92	42	87	81	83	72	79	75
10	60	69	66	87	94	71	96	61	91	89	92	84	89	86
11	73	80	78	91	97	79	98	77	96	95	98	92	95	93
12	84	82	83	97	99	83	100	83	96	99	99	95	98	96
13	96	90	92	100	99	92	—	91	96	99	100	98	99	98
14	98	96	97	—	100	96	—	97	100	100	—	99	100	99
15	98	99	99	—	—	99	—	99	—	—	—	99	—	100
16	100	99	99	—	—	99	—	100	—	—	—	100	—	100
17	—	100	100	—	—	100	—	—	—	—	—	—	—	100

NOTE: TBI = traumatic brain injury; Total Not MND = all patients including TBI not who were not malingering or suspect; No Inc = TBI no incentive; Inc Only = TBI incentive only; Not MND = all TBI not malingering or suspect; SUS = TBI suspect; MND = combined TBI probable and definite malingered neurocognitive dysfunction; CVA = cerebrovascular accident; Mem = memory disorder; Psyc = psychiatric; Total Other = all non-TBI diagnoses combined.

TABLE 6
Cumulative Percentages of Patients With Scores At or Below the Indicated Score for Maximum Span Forward Both Trials Correct (MSF2) and Maximum Span Forward (MSF1)

	All TBI					*Mild TBI*		*Moderate/Severe TBI*		*Other Diagnoses*				*Total Not*
	No Inc	*Inc Only*	*Not MND*	*SUS*	*MND*	*Not MND*	*MND*	*Not MND*	*MND*	*CVA*	*Mem*	*Psyc*	*Total Other*	*MND*
n	45	101	146	127	71	77	48	69	23	517	228	128	1,063	1,336
≤ MSF2														
2		0	0	1	3	0	4		0	2	3	0	2	1
3	0	3	2	11	24	4	29	0	13	8	11	3	8	8
4	11	12	12	29	62	17	67	6	52	34	40	20	33	30
5	40	48	45	66	87	51	92	39	78	66	73	65	68	65
6	67	70	69	84	92	73	94	65	87	86	90	87	87	85
7	87	89	88	97	99	88	100	88	96	96	96	95	96	95
8	96	99	98	99	100	99	—	97	100	99	100	100	100	99
9	100	100	100	100	—	100	—	100	—	100	—	—	—	100
≤ MSF1														
2				0	1		2			1	1		1	1
3		0	0	3	4	0	6	0	0	2	3	0	2	2
4	0	4	3	10	30	4	35	1	17	12	20	8	13	12
5	24	17	19	39	63	20	67	19	57	36	46	31	37	35
6	40	46	44	64	80	49	81	38	78	68	76	61	68	65
7	73	70	71	86	94	78	96	64	91	88	90	83	88	86
8	91	90	90	96	100	92	100	88	100	98	99	96	98	97
9	100	100	100	100	—	100	—	100	—	100	100	100	100	100

NOTE: TBI = traumatic brain injury; Total Not MND = all patients including TBI not who were not malingering or suspect; No Inc = TBI no incentive; Inc Only = TBI incentive only; Not MND = all TBI not malingering or suspect; SUS = TBI suspect; MND = combined TBI probable and definite malingered neurocognitive dysfunction; CVA = cerebrovascular accident; Mem = memory disorder; Psyc = psychiatric; Total Other = all non-TBI diagnoses combined.

TABLE 7
Cumulative Percentages of Patients With Scores At or Below the Indicated Score for Combined Maximum Forward and Backward Span (MFB)

	All TBI					*Mild TBI*		*Moderate/Severe TBI*		*Other Diagnoses*				*Total Not*
≤ MFB	*No Inc*	*Inc Only*	*Not MND*	*SUS*	*MND*	*Not MND*	*MND*	*Not MND*	*MND*	*CVA*	*Mem*	*Psyc*	*Total Other*	*MND*
n	45	101	146	127	71	77	48	69	23	517	228	128	1,063	1,336
2										1	0		1	1
3				0	0		0			1	1	0	1	1
4				1	4		6			3	4	1	3	2
5				3	6		8			6	7	1	5	4
6	0	0	0	7	11	0	17	0	0	12	12	2	10	8
7	4	2	3	14	25	3	35	3	4	21	25	9	19	17
8	18	8	11	28	47	12	54	10	30	35	44	22	34	31
9	29	13	18	48	69	18	71	17	65	57	63	42	54	49
10	36	36	36	61	80	40	81	30	78	72	76	59	69	65
11	58	53	54	76	87	57	90	51	83	85	87	75	83	79
12	67	65	66	84	94	70	96	61	91	92	96	85	91	88
13	82	79	80	93	97	86	98	74	96	96	98	90	95	94
14	84	87	86	98	99	90	98	83	100	99	100	95	98	97
15	93	95	95	99	99	96	98	93	—	99	—	96	99	99
16	96	97	97	100	100	97	100	96	—	100	—	100	100	99
17	100	100	100	—	—	100	—	100	—	—	—	—	—	100

NOTE: TBI = traumatic brain injury; Total Not MND = all patients including TBI not who were not malingering or suspect; No Inc = TBI no incentive; Inc Only = TBI incentive only; Not MND = all TBI not malingering or suspect; SUS = TBI suspect; MND = combined TBI probable and definite malingered neurocognitive dysfunction; CVA = cerebrovascular accident; Mem = memory disorder; Psyc = psychiatric; Total Other = all non-TBI diagnoses combined.

TABLE 8
Cumulative Percentages of Patients With Scores At or Below the Indicated Scaled Score for Weschler Adult Intelligence Scale—III Digit Span Scaled Score

	All TBI					*Mild TBI*		*Moderate/Severe TBI*		*Other Diagnoses*				*Total Not*
≤ Digit Span	*No Inc*	*Inc Only*	*Not MND*	*SUS*	*MND*	*Not MND*	*MND*	*Not MND*	*MND*	*CVA*	*Mem*	*Psyc*	*Total Other*	*MND*
n	27	61	88	89	59	50	39	38	20	44	42	14	140	317
2				1	2		3			2	0		1	1
3				2	7		10			2	2		2	2
4	0	0	0	6	19	0	28	0	0	7	5	0	4	3
5	8	7	7	12	36	6	41	8	25	14	10	7	10	10
6	15	8	10	20	58	10	62	11	50	16	17	7	15	15
7	30	21	24	40	76	26	80	21	70	34	31	14	29	31
8	33	36	35	54	81	36	85	34	75	43	45	29	41	43
9	41	53	49	70	88	54	90	42	85	64	60	57	63	61
10	56	64	61	79	93	68	92	53	95	73	69	86	74	72
11	63	79	74	87	97	78	97	68	95	93	79	93	86	83
12	70	85	81	93	97	84	97	76	95	96	88	100	91	89
13	82	89	86	97	97	88	97	84	95	98	88	—	92	92
14	82	97	92	99	100	96	100	87	100	100	98	—	96	96
15	89	97	94	100	—	96	—	92	—	—	100	—	98	98
16	89	97	94	—	—	96	—	92	—	—	—	—	99	98
17	96	97	97	—	—	96	—	97	—	—	—	—	99	99
18	100	98	99	—	—	98	—	100	—	—	—	—	99	99
19	—	100	100	—	—	100	—	—	—	—	—	—	100	100

NOTE: TBI = traumatic brain injury; Total Not MND = all patients including TBI not who were not malingering or suspect; No Inc = TBI no incentive; Inc Only = TBI incentive only; Not MND = all TBI not malingering or suspect; SUS = TBI suspect; MND = combined TBI probable and definite malingered neurocognitive dysfunction; CVA = cerebrovascular accident; Mem = memory disorder; Psyc = psychiatric; Total Other = all non-TBI diagnoses combined.

TABLE 9
Cumulative Percentages of Patients With Scores At or Below the Indicated Score for Maximum Span Backward Both Trials Correct (MSB2) and Maximum Span Backward (MSB1)

	All TBI					*Mild TBI*		*Moderate/Severe TBI*		*Other Diagnoses*				*Total Not*
	No Inc	*Inc Only*	*Not MND*	*SUS*	*MND*	*Not MND*	*MND*	*Not MND*	*MND*	*CVA*	*Mem*	*Psyc*	*Total Other*	*MND*
n	45	101	146	127	71	77	48	69	23	517	228	128	1,063	1,336
≤ MSB2														
0	0	0	0	2	7	0	10	0	0	13	12	1	10	8
2	13	6	8	25	42	8	46	9	35	46	47	26	41	36
3	44	37	39	66	80	39	83	39	74	75	79	57	72	68
4	71	67	69	89	97	70	98	67	96	94	95	88	93	90
5	87	84	85	96	99	88	100	81	96	98	99	93	97	96
6	96	93	94	99	100	94	—	94	100	100	100	98	99	99
7	98	99	99	100	—	99	—	99	—	—	—	99	100	100
8	100	100	100	—	—	100	—	100	—	—	—	100	—	100
≤ MSB1														
0	0	0	0	2	3	0	4	0	0	8	7	1	6	5
2	7	2	3	11	16	1	21	6	4	25	26	10	21	18
3	33	11	18	43	51	17	58	19	35	58	60	39	53	48
4	51	43	45	72	82	46	88	45	70	82	85	73	80	76
5	71	69	70	89	94	75	94	64	96	95	95	85	93	90
6	82	87	86	95	99	90	98	81	100	99	99	92	97	96
7	93	96	95	100	99	97	98	93	—	100	100	98	99	99
8	100	100	100	—	100	100	100	100	—	—	—	100	100	100

NOTE: TBI = traumatic brain injury; Total Not MND = all patients including TBI not who were not malingering or suspect; No Inc = TBI no incentive; Inc Only = TBI incentive only; Not MND = all TBI not malingering or suspect; SUS = TBI suspect; MND = combined TBI probable and definite malingered neurocognitive dysfunction; CVA = cerebrovascular accident; Mem = memory disorder; Psyc = psychiatric; Total Other = all non-TBI diagnoses combined.

TABLE 10
Percentage of Traumatic Brain Injury (TBI) Patients Scoring Below the 5th Percentile of the Total Nonmalingered Neurocognitive Dysfunction (Non-MND) Sample and of the Non-MND TBI Alone

	Cutoff	*No Incentive*	*Incentive Only*	*Suspect*	*Probable MND*	*Definite MND*
Total Non-MND sample						
RDS	≤ 4	0	0	2	9	11
MSF2	≤ 3	0	3	11	17	44
MSF1	≤ 3	0	0	3	6	0
MFB	≤ 5	0	0	3	6	6
DS	≤ 4	0	0	6	16	25
MSB2	0	0	0	2	8	6
MSB1	0	0	0	2	2	6
TBI non-MND sample						
RDS	≤ 6	2	5	16	36	50
MSF2	≤ 3	0	3	11	17	44
MSF1	≤ 4	0	4	10	30	28
MFB	≤ 7	4	2	14	28	17
DS	≤ 5	8	7	12	28	56
MSB2	≤ 2	13	6	25	38	56
MSB1	≤ 2	7	2	11	17	11

NOTE: RDS = Reliable Digit Span; MSF2 = maximum span forward both trials correct; MSF1 = maximum span forward; MFB = combined maximum forward and backward span; DS = Weschler Adult Intelligence Scale—III Digit Span scaled score; MSB2 = maximum span backward both trials correct; MSB1 = maximum span backward.

$$\text{likelihood ratio} = \text{sensitivity}/(1 - \text{specificity}) \quad (1)$$

$$\text{pretest odds} = \text{base rate}/(1 - \text{base rate}) \quad (2)$$

$$\text{posttest odds} = \text{likelihood ratio} \times \text{pretest odds}. \quad (3)$$

$$\text{PPP (posttest probability)} = \text{posttest odds}/(1 + \text{posttest odds}). \quad (4)$$

The ability to compute the PPP value oneself is important for the interpretation of the scores of individual patients. For example, when using specificity based on the performance of the entire non-MND clinical sample, a score of 5 on the RDS is associated with an overall PPP of .45 when the base rate is .30. However, the same score in a person with a TBI who is compared only to TBI patients would have a PPP of 1.00, because the false-positive error rate is 0. Of course, any scores associated with 100% specificity (false-positive error rate = 0) would be associated with a PPP of 1.00 regardless of the sensitivity as long as the sensitivity is greater than 0.

Examination of Misclassified Patients

Of the non-TBI clinical patients who scored below the 5th percentile for the RDS (≤ 6) based on the performance of the non-MND TBI sample, 54.6% were cerebrovascular accident (CVA) patients and 25.7% were memory disorder patients. Only 6 TBI patients (2.1%; 5 mild, 1 moderate/severe) scored less than 7, and all had scores of 6. Two of the TBI patients were nonnative English speakers, and another had been diagnosed with a personality disorder. The 3 remaining TBI patients, 1 of whom was the moderate/severe patient, had been described in behavioral observations as giving questionable effort. The moderate/severe patient also had a history of substance abuse, and his residential brain injury program had documented use during his inpatient stay. Overall, the non-TBI clinical patients scoring less than 7 on RDS were older (age: M = 67.8, SD = 13.9; median = 70). Age was negatively correlated with RDS in the non-MND sample ($r = -.34$; $r^2 = .12$; $p < .01$). Of all the patients scoring less than 7, 72.5% scored a 5 or a 6. Only persons with more severe neuropathology (e.g., CVA, memory disorder, brain tumor, seizure disorder) scored lower than 4, which is the 95% specificity cutoff for the entire non-MND clinical sample. In short, persons performing poorly on the RDS typically were older and had greater neuropathology. No non-MND TBI patients scored lower than 6.

DISCUSSION

The present study used a known-groups design to determine the classification accuracy of a number of Digit Span scores in the detection of cognitive malingering in TBI. In the TBI sample, more extreme scores demonstrated excellent specificity, and often, the sensitivity of the indicators was impressive even while maintaining a low false-positive error rate. The tables presented in this

TABLE 11
Sensitivity, Specificity, and Positive Predictive Power in the Traumatic Brain Injury (TBI) Sample for the Cutoffs Associated With 95% Specificity in the Total Nonmalingered Neurocognitive Dysfunction (Non-MND) Sample and in the Non-MND TBI Alone

Total non-MND sample				*Hypothetical Base Rate*				
	Cutoff	False Positive	Sensitivity	.10	.20	.30	.40	.50
RDS	≤ 4	6	10	.16	.29	.42	.53	.63
		5-8	4-19	.05-.30	.11-.49	.18-.62	.25-.72	.33-79
MSF2	≤ 3	7	24	.25	.43	.56	.67	.75
		6-9	15-36	.16-.40	.29-.60	.42-.72	.53-.80	.63-.86
MSF1	≤ 3	1	4	.31	.50	.63	.73	.80
		1-2	1-12	.05-.57	.11-.75	.18-.84	.25-.89	.33-.92
MFB	≤ 5	5	6	.14	.27	.39	.50	.60
		3-6	2-14	.04-.34	.08-.54	.13-.67	.18-.76	.25-.82
DS	≤ 4	2	19	.41	.61	.73	.81	.86
		1-5	10-31	.18-.78	.33-.89	.46-.93	.57-.95	.67-.97
MSB2	0	9	7	.09	.18	.27	.37	.47
		7-10	2-16	.02-.20	.05-.36	.08-.49	.12-.60	.17-.70
MSB1	0	5	3	.06	.13	.20	.29	.37
		4-6	1-10	.02-.22	.04-.38	.07-.52	.10-.63	.14-.71

TBI non-MND				*Hypothetical Base Rate*				
	Cutoff	False Positive	Sensitivity	.10	.20	.30	.40	.50
RDS	≤ 6	4	39	.52	.71	.81	.87	.91
		2-9	28-52	.26-.74	.44-.87	.57-.92	.67-.95	.76-.96
MSF2	≤ 3	2	24	.57	.75	.84	.89	.92
		1-6	15-36	.22-.80	.38-.90	.52-.94	.63-.96	.71-.97
MSF1	≤ 4	3	30	.53	.71	.81	.87	.91
		1-7	19-42	.23-.82	.40-.91	.54-.95	.64-.97	.73-.98
MFB	≤ 7	3	25	.48	.68	.78	.85	.89
		1-7	16-37	.20-.80	.36-.90	.49-.94	.60-.96	.70-.97
DS	≤ 5	7	36	.36	.56	.69	.77	.84
		3-14	24-49	.16-.64	.30-.80	.42-.88	.53-.92	.63-.94
MSB2	≤ 2	8	42	.37	.57	.69	.78	.84
		4-14	31-55	.20-.60	.36-.77	.49-.85	.60-.90	.69-.93
MSB1	≤ 2	3	15	.37	.57	.70	.78	.84
		1-8	8-26	.10-.74	.20-.87	.30-.92	.40-.95	.50-.96

NOTE: RDS = Reliable Digit Span; MSF2 = maximum span forward both trials correct; MSF1 = maximum span forward; MFB = combined maximum forward and backward span; DS = Weschler Adult Intelligence Scale—III Digit Span scaled score; MSB2 = maximum span backward both trials correct; MSB1 = maximum span backward.

article allow for a careful matching of a given clinical patient with the appropriate comparison group or groups. Such comparisons facilitate the interpretation of scores by helping to rule out or rule in alternative explanations for a given score. The makeup of the nonmalingering TBI groups is particularly helpful because the groups include persons with incentive. That means that the potential stress associated with a worker's compensation claim or personal injury litigation is addressed. Similarly, inclusion of stroke patients and others with neurological conditions (e.g., dementia) in the nonincentive control group helps address issues related to cognitive impairment.

Examining Tables 10 and 11 assists in determining which Digit Span scores are the most accurate in identifying malingering in TBI while minimizing false-positive errors. Table 11 compares sensitivity at approximately equal levels of specificity (95%) and also provides PPP for each of the Digit Span measures within the entire nonmalingering sample and also within only the TBI sample. For the entire nonmalingering sample, the Digit Span measures that demonstrated the best PPP are RDS, MSF2, and the Digit Span scaled score. However, when sensitivity, specificity, and PPP were investigated within only the TBI sample, all of the Digit Span scores demonstrated good sensitivity and PPP. The present study's data for Digit Span scores are consistent with several recent studies (e.g., Babikian et al., in press; Etherton, Bianchini, Greve, et al., 2005; Inman & Berry, 2002; Larrabee, 2003a; Mathias et al., 2002; Meyers & Volbrecht, 1998; Strauss et al., 2002). In particular, the data across studies are con-

sistent in demonstrating that RDS scores of less than 7 are indicative of negative response bias.

Within the TBI sample, MSB1 has the poorest sensitivity at equal levels of specificity. Within the context of the entire sample (including the neurological patients), specificity tends to be poorer at a given cutoff for scores that are computed using backward Digit Span. The difference seen between performance on forward and backward Digit Span may be related to cognitive load. Carlesimo, Fadda, Lorusso, and Caltagirone (1994) expressed this in terms of the working-memory model. They stated that forward span is a product of the articulatory loop, and backward span relies on the central executive. There is evidence of differential performance between forward and backward Digit Spans in the WAIS-III/Weschler Memory Scale—III standardization and clinical group data (Wilde & Strauss, 2002; Wilde, Strauss, & Tulsky, 2004). Ramsay and Reynolds (1995) did a meta-analysis of Digit Span tests and found that digits forward and digits backward load on different factors. There is also evidence that Digit Span forward and backward differ as a result of age, neurological damage, or different types of dementia (Carlesimo et al., 1994; Tamura, Kikuchi, Kitagawa, Otsuki, & Tashiro, 2003). In the current study, false-positive errors are more of an issue in the neurological groups for scores that involve backward Digit Span, which may be related to higher cognitive load. Because of the issue of cognitive load and the fact that Digit Span measures capacity as well as effort, comparisons should be made within a diagnostic category. Thus, the fact that the rate of positive findings with CVAs is high at a given score level does not mean that a mild TBI patient scoring at that level is likely to be a false positive; the mild TBI patient must be compared to other mild TBI patients.

It is important to emphasize that for these findings to be clinically useful, it is not necessary to identify a best or recommended cut point. The clinically relevant question is, "My patient got this score, what does this score tell me about my patient?" Given a patient's score(s), PPP can be calculated using the sensitivity and specificity data provided here or from other appropriately designed studies. For the purposes of estimating a base rate, published research (e.g., Mittenberg, Patton, Canyock, & Condit, 2002) or local base rates may be used (although from a medico-legal perspective, the clinician may prefer to used published, peer-reviewed estimates). The resultant PPP value is a percentage that describes the empirical likelihood that the observed score was produced by someone who is malingering. We have no specific recommendations as to what level of probability, but one could argue that the legal standard of more probable than not would be satisfied by a PPP of .51 or greater (Kagehiro, 1990). Reference to Table 11 shows that this standard would be met for TBI patients by RDS and both Digit Span forward scores at the 95% specificity level with at least a 30% base rate. The raw backward span cutoffs are actually at floor (it is not possible to have a maximum backward span of less than 2), so they provide a less certain basis for concluding that response bias is present except in mild TBI. Examination of the PPP data again emphasizes the importance of comparing one's own patient with the data from diagnostically similar patients, because PPP for TBI patients is lower when sensitivity and specificity are derived from the entire non-MND clinical group and the accuracy for mild TBI is better than for moderate/severe TBI. Because it is likely that the base rate of malingering in TBI is at least 20%, the oft-used convention of interpreting scores at or above the 95% specificity level (5% false-positive error rate) is reasonable, particularly in mild TBI. But again, the idea is for clinicians to calculate +PP for their specific observed scores and make their interpretations based on that.

It is important to remember that although Digit Span–based indicators provide very accurate detection of malingering in TBI, malingering detection techniques are not perfect and should not be used in isolation for the clinical diagnosis of malingering. As has been noted in many articles, manuals, and book chapters and systematized in the Slick et al. (1999) criteria, a formal diagnosis of malingering should be based on the integration of diverse clinical information. How these Digit Span variables are used within the Slick et al. system depends on the variable being examined. RDS was specifically designed to detect malingering, whereas most of the other variables, most notably the Digit Span scaled score itself, were originally designed and are still used to measure attention and working memory. The Slick et al. system treats them differently. Specifically, the B2 criterion can be met when "performance on one or more *well-validated* psychometric tests or indices designed to measure exaggeration or fabrication of cognitive deficits is consistent with feigning" (Slick et al., 1999, p. 553). RDS clearly can be applied to meet this criterion, because it was specially designed to detect malingering and is well validated. Moreover, RDS by itself is sufficient to meet Criterion B2, even in the absence of positive findings on stand-alone SVTs such as the PDRT or TOMM.

On the other hand, Criterion B6 can be met when there is "improbably poor performance on *two or more* [italics added] standardized tests of cognitive function within a specific domain (e.g., memory) that is inconsistent with documented neurological or psychiatric history" (Slick et al., 1999, p. 554). The Digit Span scaled score as well as the other span scores would contribute to meeting this criterion. It is important to recognize, however, that this criterion requires findings from two separate tests of similar

functions, so findings related to attention, concentration, or working memory, in addition to Digit Span performance, would be necessary. For example, Etherton, Bianchini, Greve, et al. (2005) presented malingering detection accuracy data for the WAIS-III Working Memory Index and its constituent subtests that could be used along with the Digit Span to meet B6. Also, the Processing Speed Index, for which similar classification accuracy data are available (Etherton, Bianchini, Heinly, & Greve, 2005), could arguably be considered an attention-concentration variable (Mapou, 1995) and could be used to meet B6 when paired with Digit Span or other attention-concentration scores validated for the detection of malingering.

We believe that these Digit Span variables should only be applied to one B criterion. The Slick et al. (1999) system allows a diagnoses of probable MND when either (a) two B criteria or (b) one B criterion and one C criterion are met. There are no applications problems relevant to Digit Span with the second case. However, there are potential problems with the first if Digit Span data are applied in a mechanistic fashion. Theoretically, a person could be diagnosed as malingering based almost exclusively on the basis of Digit Span performance with RDS meeting B2 and the Digit Span scaled score in part meeting B6. Conservative application of detection data requires that indicators that are derived in whole or in part from Digit Span performance (this would include the Working Memory Index and Mittenberg, Theroux-Fichera, Zielinski, & Heilbronner, 1995, WAIS indicators) be applied to only one B criterion. This conservative approach should also be followed within the B6 criterion. Specifically, indicators that are derived in whole or in part from Digit Span performance should apply to only one of the two improbably poor performances. Thus, positive findings on Digit Span and the Working Memory Index together would not meet B6 but on the Digit Span and Arithmetic Index would.

The current study adds to the existing literature by supplying classification accuracy for RDS, its component parts, and other Digit Span measures. The large, clinical, nonmalingering sample includes neurological patients with neuropathology and subsequent cognitive problems, TBI patients with incentive and no evidence of poor effort, and various psychiatric patients. The inclusion of these samples helps to clarify what a particular score means. If a patient with a concussion or no brain injury at all is giving good effort, his or her score should remain within the range of scores demonstrated by similar nonmalingering patients with or without incentive. The majority of neurological and psychiatric patients also did not stray far from the range of scores demonstrated by nonmalingering mild TBI scores. However, the additional cognitive demands of backward span may have been an issue for persons with significant objectively documented neuropathology (e.g., moderate/severe TBI, CVA, memory disorders, etc.), but patients with mild TBI did not appear to be similarly affected. This reinforces the idea that Digit Span scores tap into both capacity, to some degree, and effort and emphasizes the importance of interpreting a given patient's scores based on the appropriate comparison groups. When appropriately used, these Digit-Span–based scores are powerful indicators of poor effort and malingering.

References

Babikian, T., Boone, K. B., Lu, P., & Arnold, G. (in press). Sensitivity and specificity of various Digit Span scores in the detection of suspect effort. *The Clinical Neuropsychologist*.

Bernard, L. C., McGrath, M. J., & Houston, W. (1996). The differential effects of simulating malingering, closed head injury, and other CNS pathology on the Wisconsin Card Sorting Test: Support for the "pattern of performance" hypothesis. *Archives of Clinical Neuropsychology*, *11*, 231-245.

Bianchini, K. J., Mathias, C. W., & Greve, K. W. (2001). Symptom validity testing: A critical review. *The Clinical Neuropsychologist*, *15*, 19-45.

Binder, L. M. (1993). *Portland Digit Recognition Test manual* (2nd ed.). Portland, OR: Author.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegren, A., & Kaemmer, B. (1989). *MMPI-2: Manual for administration and scoring*. Minneapolis: University of Minnesota.

Butters, N., & Cermak, L. S. (1980). *Alcoholic Korsakoff's syndrome. An information-processing approach to amnesia*. New York: Academic Press.

Carlesimo, G. A., Fadda, L., Lorusso, S., & Caltagirone, C. (1994). Verbal and spatial memory spans in Alzheimer's and multi-infarct dementia. *Acta Neurologica Scandinavica*, *89*, 132-138.

Duncan, S. A., & Ausborn, D. L. (2002). The use of Reliable Digits to detect malingering in a criminal forensic pretrial population. *Assessment*, *9*, 56-61.

Etherton, J. L., Bianchini, K. J., Ciota, M. A., & Greve, K. W. (2005). Reliable Digit Span is unaffected by laboratory-induced pain: Implications for clinical use. *Assessment*, *12*, 101-106.

Etherton, J. L., Bianchini, K. J., Greve, K. W., & Heinly, M. T. (2005). Sensitivity and specificity of Reliable Digit Span in malingered pain-related disability. *Assessment*, *12*, 130-136.

Etherton, J. L., Bianchini, K. J, Heinly, M. T., & Greve, K. W. (in press). Pain, malingering, and performance on the WAIS-III Processing Speed Index. *Journal of Clinical and Experimental Neuropsychology.*

Gouvier, W. D., Hayes, J. S., & Smiroldo, B. B. (1998). The significance of base rates, test sensitivity, test specificity, and subjects knowledge of symptoms in assessing TBI sequelae and malingering. In C. R. Reynolds (Ed.), *Detection of malingering during head injury litigation* (pp. 55-80). New York: Plenum.

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesic measures with a large clinical sample. *Psychological Assessment*, *6*, 218-224.

Greiffenstein, M., Gola, T., & Baker, J. (1995). MMPI-2 validity scales versus domain specific measures in detection of factitious traumatic brain injury. *The Clinical Neuropsychologist*, *9*, 230-240.

Greve, K. W., & Bianchini, K. J. (2004). Setting empirical cutoffs on psychometric indicators of negative response bias: A methodological commentary with recommendations. *Archives of Clinical Neuropsychology*, *19*, 533-541.

Greve, K. W., Bianchini, K. J., Mathias, C. W., Houston, R. J., & Crouch, J. A. (2002). Detecting malingered performance with the Wisconsin Card Sorting Test: A preliminary investigation in traumatic brain injury. *The Clinical Neuropsychologist*, *16*, 179-191.

Heaton, S. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin Card Sorting Test manual: Revised and expanded*. Odessa, FL: Psychological Assessment Resources.

Hennekens, C. H., & Buring, J. E. (1987). *Epidemiology in medicine*. Boston: Little, Brown.

Inman, T. H., & Berry, D. T. R. (2002). Cross-validation of indicators of malingering: A comparison of nine neuropsychological tests, four tests of malingering, and behavioral observations. *Archives of Clinical Neuropsychology*, *17*, 1-23.

Iverson, G. L., & Franzen, M. D. (1994). The Recognition Memory Test, Digit Span, and Knox Cube Test as markers of malingered memory impairment. *Assessment*, *1*, 323-334.

Iverson, G. L., & Franzen, M. D. (1996). Using multiple objective memory procedures to detect simulated malingering. *Journal of Clinical and Experimental Neuropsychology*, *18*, 38-51.

Iverson, G. L., & Tulsky, D. S. (2003). Detecting malingering on the WAIS-III Unusual Digit Span performance patterns in the normal population and in clinical groups. *Archives of Clinical Neuropsychology*, *18*, 1-9.

Kagehiro, D. (1990). Defining the standard of proof in jury instructions. *Psychological Science*, *1*, 194-200.

Kiernan, R. J., Mueller, J., & Langston, J. W. (1983). *Neurobehavioral Cognitive Status Examination (Cognistat)*. Fairfax, CA: Northern California Neurobehavioral Group.

Larrabee, G. J. (2003a). Detection of malingering using atypical performance patterns on standard neuropsychological tests. *The Clinical Neuropsychologist*, *17*, 410-425.

Larrabee, G. J. (2003b). Exaggerated MMPI-2 symptom report in personal injury litigants with malingered neurocognitive deficit. *Archives of Clinical Neuropsychology*, *18*, 673-676.

Larrabee, G. J. (Ed.). (2005). Assessment of malingering. In *Forensic neuropsychology: A scientific approach* (pp. 115-158). New York: Oxford University Press.

Mapou, R. L. (1995). A cognitive framework for neuropsychological assessment. In R. L. Mapou & J. Spector (Eds.), *Clinical neuropsychological assessment: A cognitive approach* (pp. 295-337). New York: Plenum.

Mathias, C. W., Greve, K. W., Bianchini, K. B., Houston, R. J., & Crouch, J. A. (2002). Detecting malingered neurocognitive dysfunction using the Reliable Digit Span in traumatic brain injury. *Assessment*, *9*, 301-308.

Meyers, J. E., & Diep, A. (2000). Assessment of malingering in chronic pain patients using neuropsychological tests. *Applied Neuropsychology*, *7*, 133-139.

Meyers, J. E., & Volbrecht, M. (1998). Validation of Reliable Digits for detection of malingering. *Assessment*, *5*, 303-307.

Mild Traumatic Brain Injury Committee of the Head Injury Interdisciplinary Special Interest Group of the American Congress of Rehabilitation Medicine. (1993). Definition of mild traumatic brain injury. *Journal of Head Trauma Rehabilitation*, *8*, 86-87.

Millis, S. R. (2004). Evaluation of malingered neurocognitive disorders. In M. Rizzo & P. J. Esslinger (Eds.), *Principles and practice of behavioral neurology and neuropsychology* (pp. 1077-1089). Philadelphia: W. B. Saunders.

Millis, S. R., Putnam, S. H., Adams, K. M., & Ricker, J. H. (1995). The California Verbal Learning Test in the detection of incomplete effort in neuropsychological evaluation. *Psychological Assessment*, *7*, 463-471.

Millis, S. R., & Volinsky, C. T. (2001). Assessment of response bias in mild head injury: Beyond malingering tests. *Journal of Clinical and Experimental Neuropsychology*, *23*, 809-828.

Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Baserates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology*, *24*, 1094-1102.

Mittenberg, W., Theroux, S., Aguila-Puentes, G., Bianchini, K., Greve, K., & Rayls, K. (2001). Identification of malingered head injury on the Wechsler Adult Intelligence Scale—3rd edition. *The Clinical Neuropsychologist*, *15*, 440-445.

Mittenberg, W., Theroux-Fichera, S., Zielinski, R. E., & Heilbronner, R. L. (1995). Identification of malingered head injury on the Wechsler Adult Intelligence Scale—Revised. *Professional Psychology: Research and Practice*, *26*, 491-498.

Ramsay, M. C., & Reynolds, C. R. (1995). Separate digits tests: A brief history, a literature review, and a reexamination of the factor structure of the Test of Memory and Learning (TOMAL). *Neuropsychology Review*, *5*, 151-171.

Rogers, R. (Ed.). (1997). Researching dissimulation. In *Clinical assessment of malingering and deception* (pp. 398-426). New York: Guilford.

Slick, D. J., Sherman, E. M. S., & Iverson, G. L. (1999). Diagnostic criteria for malingering neurocognitive dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist*, *13*, 545-561.

Stern, R. A., & White, T. (2003). *The Neuropsychological Assessment Battery*. Lutz, FL: Psychological Assessment Resources.

Strauss, E., Slick, D. J., Levy-Bencheton, J., Hunter, M., MacDonald, S. W. S., & Hultsch, D. F. (2002). Intraindividual variability as an indicator of malingering in head injury. *Archives of Clinical Neuropsychology*, *17*, 423-444.

Suhr, J. A., & Boyer, D. (1999). Use of the Wisconsin Card Sorting Test in the detection of malingering in student simulator and patient samples. *Journal of Clinical and Experimental Neuropsychology*, *21*, 701-708.

Tamura, I., Kikuchi, S., Kitagawa, M., Otsuki, M., & Tashiro, K. (2003). Age-related working memory decline in patients with Parkinson's disease. *No To Shinkei*, *55*, 335-340.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *The Stanford-Binet Intelligence Scale, 4th ed.* Chicago: Riverside.

Tombaugh, T. (1996). *Test of Memory Malingering manual*. New York: MultiHealth Systems.

Wechsler, D. A. (1981). *Wechsler Adult Intelligence Scale—Revised*. New York: Psychological Corporation.

Wechsler, D. A. (1987). *Wechsler Memory Scale—Revised manual*. New York: Psychological Corporation.

Wechsler, D. A. (1997a). *Wechsler Adult Intelligence Scale—III*. New York: Psychological Corporation.

Wechsler, D. A. (1997b). *Wechsler Memory Scale—III*. New York: Psychological Corporation.

Wilde, N., & Strauss, E. (2002). Functional equivalence of WAIS-III/WMS-III digit and spatial span under forward and backward recall conditions. *The Clinical Neuropsychologist*, *16*, 322-330.

Wilde, N. J., Strauss, E., & Tulsky, D. S. (2004). Memory span on the Wechsler Scales. *Journal of Clinical and Experimental Neuropsychology*, *26*, 539-549.

Matthew T. Heinly, M. S., is a doctoral student in applied biopsychology at the University of New Orleans, Louisiana. His current research interests include cognitive neuropsychology, brain trauma, and malingering.

Kevin W. Greve, Ph.D., is Professor of Psychology at the University of New Orleans and in private practice as a board certified clinical neuropsychologist in the New Orleans area. His current research interests include executive functions, aggression, and response bias/malingering.

Kevin J. Bianchini, Ph.D., is a board certified clinical neuropsychologist in private practice in New Orleans. His current research interests include pain-related disability, malingering, and toxic exposure.